WHAT IS CLAIMED IS:

1. A method of updating a personalized web page, the method comprising:

identifying characteristic features and an information sample from an input that specifies a web site, the web site providing the information sample;

determining corresponding contents of the web site based on the characteristic features;

extracting the corresponding contents on the basis of relevancy of the corresponding contents to the information sample; and

updating the personalized web page with the corresponding contents.

2. The method of claim 1 further comprising repeating the steps of determining, extracting, and updating at a frequency specified by the input.

3. The method of claim 2 further comprising:

requesting a verification after updating the personalized web page;

if the verification confirms the update, adding the corresponding contents into a training set; and

locating the corresponding contents of the web site based on the training set in the repeating steps.

4. The method of claim 1 wherein identifying characteristic features include identifying a topic keyword.

5. The method of claim 1 wherein identifying the characteristic features include identifying a layout of the information sample.

6. The method of claim 1 wherein identifying the characteristic features include identifying a domain keyword in the information sample.

7. The method of claim 1 wherein identifying the characteristic features include identifying a semantic category in the information sample.

14

8.   The method of claim 1 wherein identifying the characteristic features include identifying an event in the input.

9.   The method of claim 1 wherein determining the corresponding contents further includes:

assigning a score according to degree of the relevancy of the corresponding contents to the information sample; and

requesting a different input if the score is below a pre-determined threshold.

10.   A system for updating a personalized web page, the system comprising:

a processor, which is connected to a plurality of web sites by a network, the processor being adapted to:
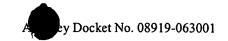
identify characteristic features and an information sample from an input that specifies a web site among the plurality of web sites, the specified web site providing the information sample;

determine corresponding contents of the web site based on the characteristic features;

extract the corresponding contents on the basis of relevancy of the corresponding contents to the information sample; and

update the personalized web page with the corresponding contents; and

a display, which displays the personalized web page.

11.   The system of claim 10 wherein the processor is adapted to repeatedly identify, determine, and update at a frequency specified by the input.

12.   The system of claim 10 wherein the processor is further adapted to:

request a verification after updating the personalized web page;

if the verification confirms the update, add the corresponding contents into a training set; and

locate the corresponding contents of the web site based on the training set in the repeating steps.

13.   The system of claim 10 further comprising a topic database for the processor to identify one of the characteristic features as a topic keyword.

14.   The system of claim 10 further comprising a layout analyzer for the processor to identify one of the characteristic features as a layout.

15.   The system of claim 10 further comprising an online domain database for the processor to identify one of the characteristic features as a domain keyword.

16.   The system of claim 10 further comprising a semantic tree for the processor to identify one of the characteristic features as a semantic category.

17.   The system of claim 10 further comprising an event database for the processor to identify one of the characteristic features as an event.

18.   The system of claim 10 wherein the processor is further adapted to:

assign a score based on degree of the relevancy of the corresponding contents to the information sample; and
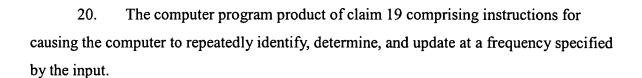
request a different input if the score is below a pre-determined threshold.

19.   A computer program product residing on a computer readable medium comprising instructions for causing a computer to:

identify characteristic features and an information sample from an input that specifies a web site, the web site providing the information sample;

determine corresponding contents of the web site based on the characteristic features;

extract the corresponding contents on the basis of relevancy of the corresponding contents to the information sample; and

update the personalized web page with the corresponding contents.

20.   The computer program product of claim 19 comprising instructions for causing the computer to repeatedly identify, determine, and update at a frequency specified by the input.

21.   The computer program product of claim 19 comprising instructions for causing the computer to:

request a verification after updating the personalized web page;

if the verification confirms the update, add the corresponding contents into a training set; and

locate the corresponding contents of the web site based on the training set in the repeating steps.

22.   The computer program product of claim 19 wherein the characteristic features include a topic keyword, a layout, a domain keyword, a semantic category, and an event.

23.   The computer program product of claim 22 comprising instructions for causing the computer to:

assign a score based on degree of the relevancy of the corresponding contents to the information sample; and
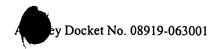
request a different input if the score is below a pre-determined threshold.

24.   The computer program product of claim 23 comprising instructions for causing the computer to:

compare the topic keyword and the layout of the corresponding contents with those of the information sample to determine the degree of relevancy.

25.   The computer program product of claim 23 comprising instructions for causing the computer to:

compare the domain keyword, the semantic category, the event, and the layout of the corresponding contents with those of the information sample to determine the degree of relevancy.